Inexpensive and easy method for 6 fragment Golden Gate Assembly of a modular S/MARs mammalian expression vector and its variants

Authors: IONUT ADRIAN PASCU[1*]
1 Faculty of Biology, University of Bucharest, Bucharest, Romania
*Author to whom correspondence should be addressed: E-mail: ionut.cepleanu@gmail.com

Abstract

A basic requirement for synthetic biology is the availability of efficient DNA assembly methods. Numerous methods have been previously reported to accomplish this task. One such method has been reported, which allows parallel assembly of multiple DNA fragments in a one-tube reaction, called Golden Gate Assembly. Here described is a simplified and inexpensive Golden Gate Cloning Protocol in which the amplified PCR fragments that enter the one-step-one-pot reaction are stored in Zymo DNA/RNA Shield at -20 degrees C and thawed whenever needed to be used as fragments or modules in the assembly. The protocol inludes the design step, in which fragments are designed to posses unique overhangs, amplification of modules using a high fidelity polimerase from preexisting plasmids or gene fragments, clean-up of the PCR products (fragments) in one tube, assembly, DpnI digestion for eliminating the background plasmids that remain after the PCR reaction and transformation of the resulting assembly reaction into competent *E.coli* cells. The protocol eliminates the need for vectors and inserts. The plasmid is build using only PCR products or fragments, with one of them containing the bacterial origin of replication and the antibiotic resitance gene for selection. Multiple modular plasmids can be constructed in just one day with minimal hands-on time.

Key words: ***Golden Gate, Cloning, S/MARS, mammalian expression vector***

Introduction

The principle of Golden Gate Cloning[1] is based on the ability of type IIS enzymes to cleave outside of their recognition site, allowing two DNA fragments flanked by compatible restriction sites to be digested and ligated seamlessly. Since the ligated product of interest does not contain the original type IIS recognition site, it will not be subject to redigestion in a restriction-ligation reaction.

All products that reconstitute the original site will be re-digested, allowing their components to be made available for further ligation, leading to formation of an increasing amount of the desired product with increasing time and cycles of incubation. Since the sequence of the overhangs at the ends of the digested fragments can be chosen to be any 4-nucleotide sequence, multiple compatible DNA fragments can be assembled in a defined linear order in a single restriction-ligation step.

Based on this method multiple standards to construct multi-gene circuits have been developed. Mo-Clo, or the modular cloning standard, is one of these standards [2]. Mo-Clo consists of standard parts, or basic modules, of eukariotic Promoters, 5'UTR's, Signal Peptides, CDS's, 3'UTR's and Terminator regions, each module containing a standardized 4 nucleotide overhang. Mo-clo or similar standards like Golden Braid 2.0 [3] use 3 levels of succesive assemblies.

Level 0, or basic modules (Promoters, 5'UTR's, Signal Peptides, CDC's, 3'UTR's and Terminator regions) are assembled into Level 1 transcriptional units using a Type IIs restriction enzyme. The level 1 transcriptional units can be, in turn, ligated into level 2 multigene constructs using a different type IIs restriction enzyme.

However, although standardized and agreed by the community, these standards may proove too complex for assembling just a few/several fragments or for students that are just getting started in molecular biology or that have a desire for more customisation.

In turn, here described is a simple inexpensive protocol for standardizing and assembling several fragments with good fidelity, in just one day with minimal *hands-on* time.

For this purpose a modular S/MARs mammalian expression vector (called pNoname) has been designed. A S/MARs vector is a type of non-integrating episomal vector that is used for mammalian expression[4][5]. It contains a region called a Scaffold/Matrix Attachment Region which is an element designed to permit the persistance of the vector as an extrachromosomal element within a host cell [6]. It also provides the vector with the ability to remain episomal and be mitotically stable during cell division [5]. Furthermore, these types of vectors can replicate during cell division in certain types of cells. [7].

The modular vector designed in this paper contains 6 fragments or modules that can be easily interchangeable.

In order to facilitate the Golden Gate Assembly, fragments to enter the assembly, previously amplified by PCR can be stored in Zymo DNA/RNA Shield at a ratio of 1:3 (PCR :

Zymo Shield) thus eliminating the need to store the fragments in plasmids and facilitating the screening of the desired mutants.

Zymo DNA/RNA shield protects DNA from freeze thaw cycles, stabilizes nucleic acids and deactivates nucleases. [8]. DNA overhangs of several bases designed in primers can further protect the PCR products from exonucleases thus leaving the type IIs restriction sytes intact.

After designing the primers to posess specific 4 base pair overhangs and accomodating the DNA fragments that will enter the assembly reaction to lack any other specific type IIs restriction sytes (depending on the type IIs restriction enzyme used) [9] the fragments can be cleaned , assembled and the resulting product transformed into competent *E.coli* cells (*home-made* or purchased) to generate the desired mutants.

Multiple plasmids can be constructed at the same time providing that certain rules are folowed and provided that the correct fragments are mixed for each reaction.

Fragments assemblies (called F or modules) and transformations can be performed in the same day. The core fragment which contains the bacterial origin of replication and the antibiotic gene can be amplified by PCR. Multiple types of core fragments could be used, with interchangeable antibiotic resistance genes. Of the amplified PCR products stored in ZS (Zymo Shield) 20 µl can be added to the assembly reaction (of each fragment). Before assembly, the PCR mix (mixed fragments) is cleaned and concentrated and 20 µl of eluate is used in the GG assembly reaction. Multiple reactions can be made at the same time. After the assembly, DpnI digestion is performed to remove the background plasmid that was used as a PCR template and then 2-5 µl is transformed into competent E.coli.

Materials and Methods

pNoname design and construction

In order to provide a backbone for future experiments a modular S/MARs vector called pNoname was designed. The vector consists of 6 modules, designed to be easily interchangeable in a short amount of time and with minimum financial effort.

1) One module consists of the required **bacterial origin of replication** for replication and maintenance in bacteria **and the antibiotic resistance** gene for selection.
2) One module consists of the **promoter and enhancer region**.
3) One module consists of the **reporter gene (cGFP)** with a kozack consensus sequence.[10]

4) One module consists of a **shortened Nuclear Matrix Attachment Region**, MARS5 [7] directly downstream of the gene.
5) One module consists of the **Terminator** and the poly adenylation sequence.
6) One module consists of the **SV40** (Simian Virus 40) **DTS** (DNA nuclear tergetting sequence), a 72 bp repeat of the early enhancer, that has been shown to mediate nuclear transport of the inserted DNA.

The primers for amplifiying the fragments or modules were designed using the New England Biolabs Golden Gate Assembly Tool (https://goldengate.neb.com/#!/) in order to generate specific 4 base pairs overhangs and were further modified to accomodate other requirements (GC content, TM etc). The standardized 4 base pairs overhangs are highlighted in Fig. 1. After digestion with BsaI and ligation with T4 ligase the BsaI recognition sequence is eliminated from the final construct thus preventing redigestion. The primers used for the amplification of the modules are shown in Fig. 3.

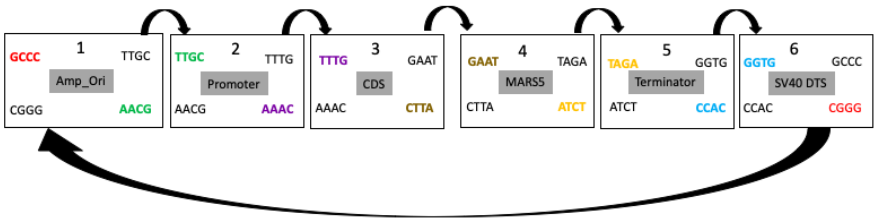

Fig. 1 Overhangs generated for each module with the assemble direction given in arrows.

A schematic representation of pNoname variant 1 (pNoname_CMV_GFP_BGH) is shown in Fig.2.

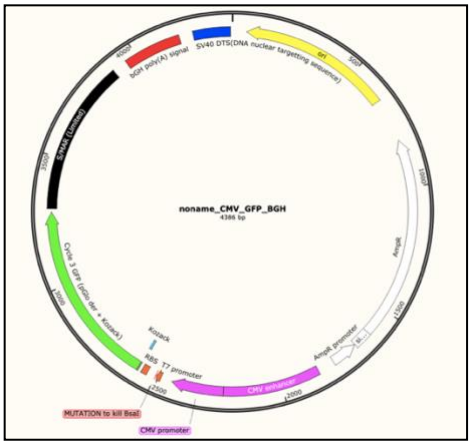

Fig.2 pNoname variant CMV_GFP_BGH after assembly that contains all 6 modules (Promoter region, Coding sequance, MARS5, Terminator and poly adenilation signal, SV40 DTS and the core fragment (bacterial origin of replication and antibiotic resistance gene)

Domestication (removal of internal BsaI sites)

In order to enter the Golden Gate Assembly, all other BsaI sites except the ones containing the overhangs must be eliminated from the fragments.

Domestication of the plasmids was performed using primers FPCTokillmyBsaI, RPCTokillmyBsaI, FTokillBsaI, RtokillBsaI with a modified version of Site Directed Mutagenesis, called SPRIP (Single Primer Reaction in Parallel) [11]. Two plasmids were domesticated, pUC18 and pcDNA3.1(+) (ThermoFischer Cat.No V79020). The primers were designed to be complementary to each other and to contain a single base pair mutation that will eliminate the BsaI from the ampicilin resistance of the pUC18 plasmid and eliminate the BsaI site just upstream of the T7 promoter of pcDNA3.1(+) respectively. Single primer PCR was performed in 2 different tubes for each plasmid using Q5 2X Master Mix(New England Biolabs Cat.No M0492S). After the amplification, the PCR tubes contents were mixed and allow to re-anneal using the following cycle (5 min at 95° C; 1 min at 90° C; 1 min at 80°C, 0.5 min at 70°C, 0.5 min at 60°C, 0.5 min at 50°C, 0.5 min at 40°C, hold at 37°C). After the re-annealing step, DpnI (New England Biolabs Cat.No R0176S) digestion was performed prior to transformation to eliminate the parental plasmid. Transformed cells (New England Biolabs *E.coli* Turbo Cat.No) made competent using the protocol in Supplemental Materials were placed in the incubator overnight at 37° and colonies were screened for desired mutants using BsaI-HF®v2 (New England Biolabs Cat.No R3733S).

Construction of modules (fragments) for assembly

The Amp_Ori region was amplified from pUC18 after it's *domestication* (removal of any internal BsaI sites). The other modules used in this experiment , CMV promoter and Enhancer, the

Primer Name		Sequence 5' - 3'	Length	Overhang
Amp+Ori_fwd_pUc18	5'-3'	ggctac**ggtctca***gccc*ctaactccAGCAAAAGGCCAGGAACC	43	*GCCC*
Amp+Ori_rev_pcUc18	5'-3'	ggctac**ggtctct***gcaa*GTCGACGTCAGGTGGCAC	35	gcaa
Promoter__CMV_fwd	5'-3'	ggctac**ggtctct***ttgc*TTCGCGATGTACGGG	32	ttgc
Promoter_CMV_rev	5'-3'	ggctac**ggtctca***caaa*attatttcAAGTTTAAACGCTAGCCAGCTTG	48	caaa
GFPKozack_fwd	5'-3'	ggctac**ggtctca***TTTG*TTTAACTTTAAGAAGGAGATATAC	41	tttg
GFPKozack_rev	5'-3'	ggctac**ggtctcc***ATTC*ATTATTTGTAGAGCTCATC	36	attc
MARS5_fwd	5'-3'	ggctac**ggtctcc***gaat*tcgagctcCTGTGCTGTCGCCTAACTATC	46	gaat
MARS5_rev	5'-3'	ggctac**ggtctca***tcta*gactcgagTCATGGGATCCAAGTATCAAG	46	tcta
TERM_BGH_fwda	5'-3'	ggctac**ggtctca***TAGA*GGGCCCGTTTAAACC	32	taga
TERM__BGH_rev	5'-3'	ggctac**ggtctcg***ccac*agaatAGCCCCAGCTGGTTCTTTC	41	ccac
DTS40_fwd	5'-3'	ggctac**ggtctcg***GTGG*AATGTGTGTCAGTTAGGG	35	gtgg
DTS40_rev	5'-3'	ggctac**ggtctca***GGGC*GGGACTATGGTTGC	31	***gggc***
TERM_Sv40_fwd	5'-3'	*ggctac**ggtctca**TAGA CTCATGCTGGAGTTCTTCG*	*35*	*taga*
TERM_Sv40_Rev	5'-3'	*ggctac**ggtctcg**ccac TGACCATGATTACGCCAAGC*	*37*	*ccac*
PROMO_Sv40_Fwd	5'-3'	*ggctac**ggtctcg**TTGC AATGTGTGTCAGTTAGGG*	*35*	*ttgc*
PROMO_Sv40_Rev	5'-3'	*ggctac**ggtctca**caaa ACCTGTCTCTTGATCAGATC*	*37*	*caaa*

Fig. 3 Primers used in building the fragments for GGA (Golden Gate Assembly). In bold the BsaI restriction syte contained in each primer. The 4 base pair overhang providing direction in assembly is shown in italic. In capital letters are the regions of the primers that anneal to the template.

SV40 promoter, the SV40-DTS, the BGH terminator region and the SV40 terminator region were amplified from pcDNA3.1(+) after domestication.

Gfp_Kozack (The reporter gene coding for a green fluorescent protein) and MARS5 were synthetized by Twist Biosciences. Primers were used to further multiply the sequences from pNoname.

All amplifications were made using a high fidelity polymerase (New England Biolabs Q5 2X Master Mix) according to the manufacturers recommandations in a 25 µl final reaction volume.

Construction of pNoname variants

After amplification, 20µl of the PCR product that contained the specific modules were stored in Zymo DNA/RNA shield (Zymo Research Cat.No R1100-50) in the freezer at -20 degrees Celsius. 5µl of the PCR product was loaded on a gel for confirmation of the reaction.

20 µl of each module (6 in total) were mixed in a single tube and cleaned using the Zymo Research DNA Clean and Concentrator -5 kit (Cat.No D4013). Each tube contained fragments from a single pNoname variant. A total of 4 pNoname variants were constructed. (pNoname_CMV_GFP_BGH, pNoname_SV40_GFP_SV40, pNoname_CMV_GFP_SV40, pNoname_SV40_GFP_BGH). The assembly reaction was set up as follows. 20 µl of mixed purified fragments (6 fragments in total) , 1.5 µl (600 U) of T4 ligase (New England Biolabs Cat.No M0202S), 1 µl (20 U) of BsaI-HF®v2, 2,5 µl of T4 Ligase Buffer. 5 µl of purified fragments were kept for gel agarose confirmation.

The assembly cycle was set up as follows (1 min at 16° C + 1 min at 37°C) X 30 cycles followed by a 5 minute soak time at 60°C. DpnI digestion was performed to eliminate any background plasmids from the PCR reaction using 1 µl of DpnI (37°C for 15 minutes followed by 5 minutes at 60°C) and transformed with 5µl of assembly product using *E.coli* Neb Turbo made competent by the transformation protocol in Supplemental Materials.

Cells were plated on antibiotic selection plates with LB Agar containing 100 µg/ml ampicilin and left in the incubator overnight at 37°C.

Mutants were confirmed by restriction digest on a 1% agarose gel and by amplifying the regions using primers provided in Fig.3.

The detailed design and assembly protocol is provided in Supplemental Materials.

Results and discussions

Domestication (removal of internal BsaI sites)

Domestication of the two plasmids used (pUC18 and pcDNA31(+)) was performed using Single Primer Reaction in Parallel. Mutants were screened by restriction digest with BsaI to confirm the succesful removal of the BsaI restriction site. After the restriction digest BsaI should not be able to linearise the pUC18-BsaI mutant, thus, compared to the wild type variant the mutant should migrate as the supercoiled form in the gel. pcDNA3.1(+) contains two BsaI restriction sytes. The mutant form, pcDNA3.1(+)-BsaI, after domestication, should migrate as one linearised fragment of 5428 base pairs on the gel, as opposed to the wild type variant that should generate 2 fragments.

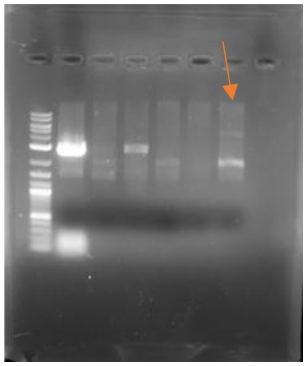

Fig. 4 Gel electrophoresis after BsaI restriction digest. Lane 1 Ladder, Lane 2 wild type pUC18 (Control), lanes 3,4,5,6,7 screened colonies after miniprep digested with BsaI

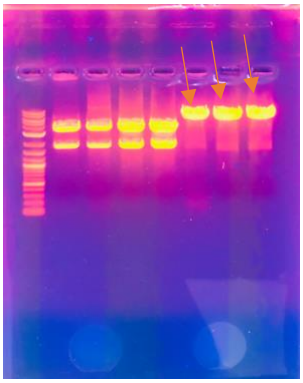

Fig. 5 Gel electrophoresis after BsaI restriction digest. Lane 1 Ladder, Lane 2 wild type pcDNA3.1(+) (Control), lanes 3,4,5,6,7,8 screened colonies after miniprep digested with BsaI.

Construction of modules (fragments) for assembly

The fragments (modules) needed for assembly were amplified using a high fidelity polymerase in a final reaction volume of 25 µl, according to the suppliers recommandations. 5 µl of each reaction was loaded onto a 1% agarose gel next do a DNA ladder to confirm succesful amplification. 20 µl of each reaction was stored at a ratio of 1:3 in Zymo DNA/RNA Shield and aftert proper labelling, the tubes were placed in the freezer at -20 °.

Construction of pNoname variants

Four pNoname variants were constructed in the same day using 20µl of each module previously amplified and stored in Zymo Shield. The six modules were loaded in each of the 4 tubes coressponding to each variant to assemble. A gel electrophoresis pre and post assembly confirmed the succesful assembly of the variants.

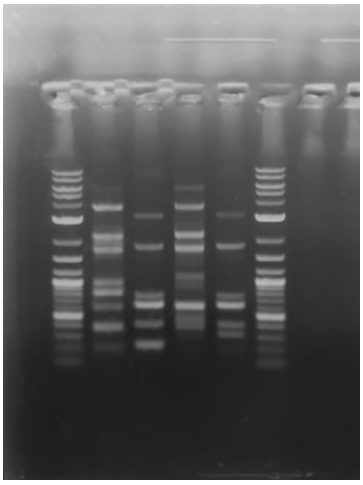

Fig. 6 Lane 1 Ladder, Lane 2 pNoname_SV40_GFP_SV40 post assembly, Lane 3 pNoname_SV40_GFP_SV40 pre assembly, Lane 4 pNoname_SV40_GFP_BGH post assembly, pNoname_SV40_GFP_BGH pre assembly

After transformation 2-4 collonies from each of the four plates were harvested for screening.

Desired mutants were confirmed by restriction digest according to their individual DNA sequences. Out of **2 mutants** screened for **pNoname_SV40_GFP_SV40** <u>*both*</u> showed the restriction pattern consistent with a correct assembly. Out of **2 mutants screened** for **pNoname_SV40_GFP_BGH** <u>*one*</u> showed the desired pattern. For **pNoname_CMV_GFP_BGH** *3* **out of 4** colonies screened showed the desired pattern and *1* out of **2 collonies** screened showed the desired pattern for **CMV_GFP_SV40 respectively** (data not shown). Although few colonies were screened, mutants were confirmed in all that were processed.

Fig. 6 pNoname variants. Lane 1 Ladder 1kb plus, Lane 2 pNoname_SV40_GFP_SV40 digested with XhoI, Lane 3 pNoname_SV40_GFP_BGH digested with XhoI, Lane 4 pNoname_SV40_GFP_SV40 digested with AseI, Lane 5 pNoname_SV40_GFP_BGH digested with AseI, Lane 6 pNoname_SV40_GFP_SV40 non-digested, Lane 7 pNoname_SV40_GFP_BGH non-digested

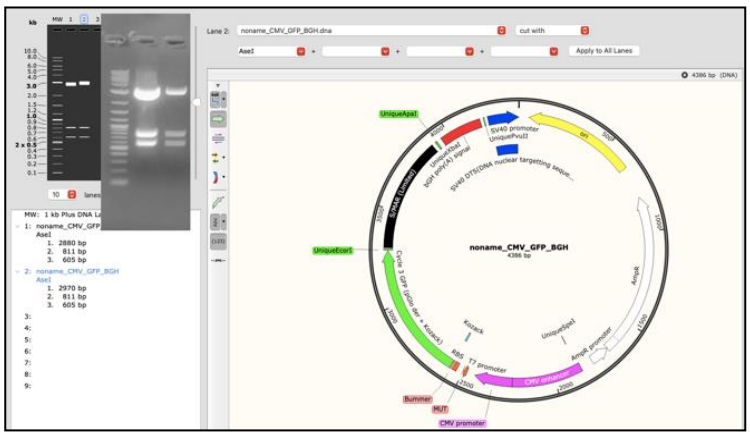

Fig. 8 pNoname variants. Lane 1 Ladder 1kb plus, Lane 2 pNoname_CMV_GFP_SV40 digested with AseI, Lane 3 pNoname_CMV_GFP_BGH digested with AseI

The number of colonies obtained after the transformation with the assembled DNA was low, for all 4 plasmid variants, especially considering the fact that the whole outgrowth after the 1 h incubation step at 37 °C was concentrated (outgrowth was spinned at 1000G for 6 minutes and the bacterial pellet resuspended in 100 µl of LB Broth and plated; For the transformation procedure see Transformation Protocol 1 in Supplementary Materials).

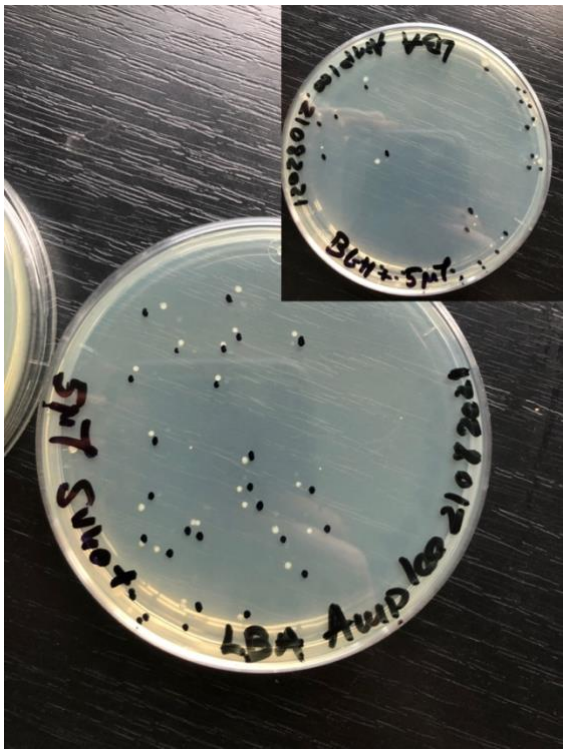

Fig. 7 Colonies generated after transforming E.coli Neb Turbo with 5 µl of assembled pNoname_SV40_GFP_SV40 and 5 µl of pNoname_SV40_GFP_BGH. 27 and 25 colonies per plate respectively

For example, 5 µl of transformed pNoname_SV40_GFP_SV40 generated about 25 colonies per plate after transformation, 5 µl of pNoname_SV40_GFP_BGH generated 28 colonies after transformation into *home-made* competent *E.coli*. Similarly, pNoname_CMV_GFP_SV40 and pNoname_CMV_GFP_BGH generated aproximately the same number of colonies after transformation with 5µl of assembled product. (Data not shown). However, this fact (low transformation efficiency) did not present itself as an inconvenience due to the fact that assembly efficiency was high, considering that the ratio of the correctly assembled plasmids to incorrect ones was high.

pNoname_CMV_GFP_BGH fluorescent expression testing.

The pNoname variants CMV_GFP_BGH and CMV_GFP_SV40 (the variants that contain the CMV enhancer and promoter regions) were designed to contain a T7 promoter for ease of screening in bacteria prior to transfection in mammalian cells.

The T7 promoter is a sequence of DNA 18 base pairs long up to transcription start site at +1 (5' – TAATACGACTCACTATAG – 3') that is recognized by T7 RNA polymerase [12]. It's induction with IPTG is possible when the vector that contains it is expressed in a bacterial strain specifically engineered to code for T7 RNA polymerase, such as *E.coli* Bl21(DE3). *E.coli* Bl21(DE3) posesses the lacUV5 promoter that expresses the T7 RNA polymerase upon induction with IPTG. Results show that induction of pNoname_CMV_GFP_BGH promotes GFP expression when transformed in BL21(DE3) compared to it's non-IPTG induced counterpart, albeit the expression is not a strong one, no doubt also limited by the lack of a proper bacterial terminator.

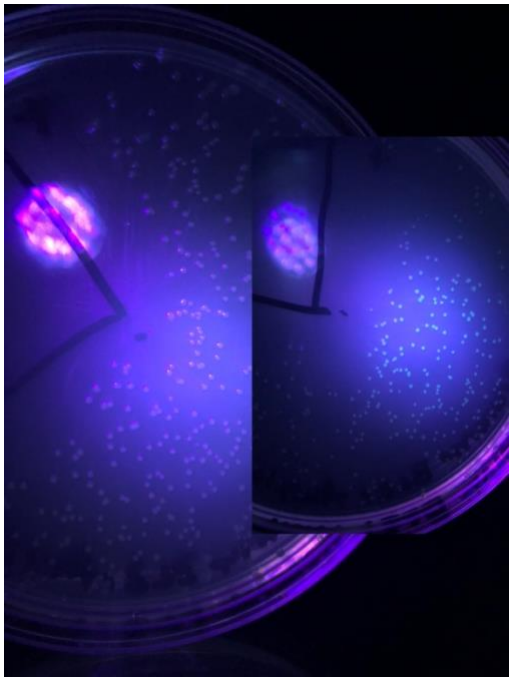

Fig. 9 pNoname_CMV_GFP_BGH transfromed in E.coli BL21(DE3). Induced vs. Non-induced with IPTG under UV flashlight

When transfected in mammalian HEK293T with PEI the EGFP expression is present although no further experiments beyond screening for mammalian expression using a fluorescent microscope have been made at this point.

The ability of the vector to be maintaned as an episome inside HEK293 cells, the transfection efficiency, the intensity of the fluorescence have not been evaluated. The cells have not been passaged and the fluorescence generated by the cGFP gene was not monitored beyond the 120 hour mark.

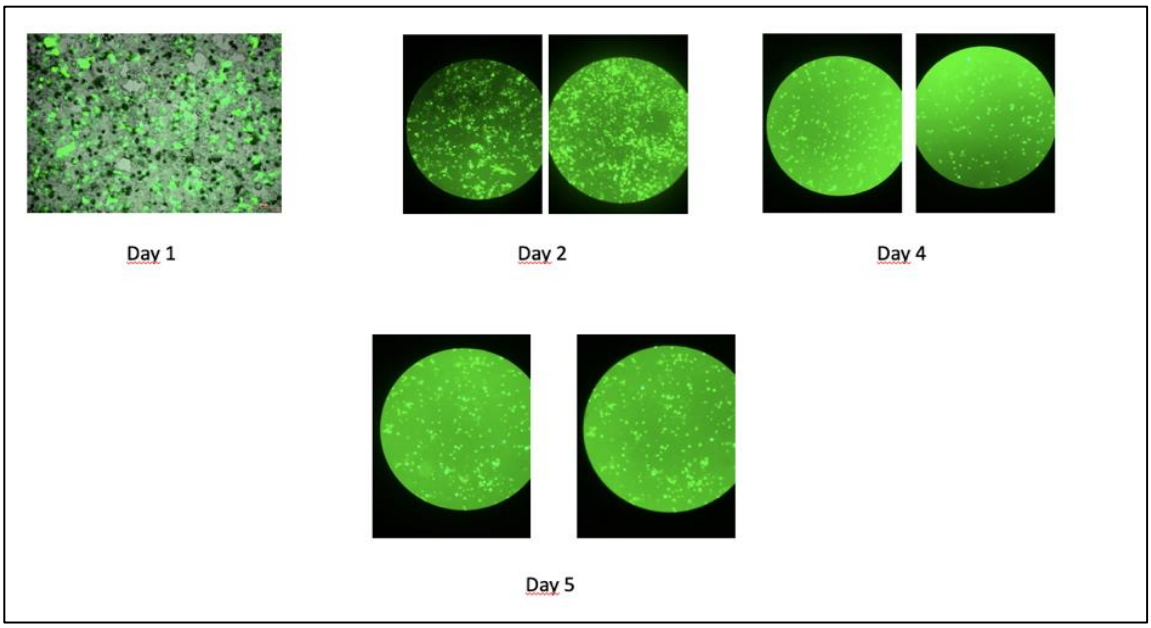

Fig. 9 HEK293T cells transfected with pNoname_CMV_GFP_BGH using PEI. Expression was monitored using a fluorescence microscope at day 1, day 2, day 4 and day 5 post transfection.

CONCLUSIONS

The need for cheap, easy methods to design and build various expression vectors is still present in molecular and synthetic biology, especially for small laboratories that lack the capacity to implement large scale screening and/or construction of many vectors at a time and complicated gene circuits. By careful design and with cost savings in mind, small laboratories and students can build vectors without the need for multiple restriction enzymes, laborious screening and using just a few selected products.

The protocol presented here could be such an alternative. Counterintuitively, by building the plasmids from the ground up every time instead of cloning a single gene, using stored modules amplified by PCR (including the core element consisting of the bacterial origin of replication and the antibiotic resistance gene), instead of using the concept of vectors and inserts, the workflow could be simplified.

Storing the amplified PCR fragments or modules in an environment that is endo-and exonuclease free, protecting the nucleic acids from freeze thaw cycles, amplifying further modules with a high fidelity polymerase that has proof reading capabilities when needed using just 1µl of parental plasmid, cleaning the PCR products just one time per assembly reaction, customizing the modules through careful design and using just one restriction enzyme for molecular cloning could

be a reliable and cheap method for multiple modular vector construction in a very short time, reducing hand-on time and costs.

Acknowledgements:

I would like to express my very great appreciation to Sebastian.S Cocioba for providing guidance and the bacterial strains used in this paper, to Lector Dr. Ionela Sarbu for providing the primers used and to Dr. Rodica-Aura Badea for the help provided in testing pNoname expression in HEK293 cells.

REFERENCE

1. Engler C, Kandzia R, Marillonnet S (2008) A One Pot, One Step, Precision Cloning Method with High Throughput Capability. PLoS ONE 3:e3647. https://doi.org/10.1371/journal.pone.0003647

2. Weber E, Engler C, Gruetzner R, et al (2011) A Modular Cloning System for Standardized Assembly of Multigene Constructs. PLOS ONE 6:e16765. https://doi.org/10.1371/journal.pone.0016765

3. Sarrion-Perdigones A, Falconi EE, Zandalinas SI, et al (2011) GoldenBraid: An Iterative Cloning System for Standardized Assembly of Reusable Genetic Modules. PLOS ONE 6:e21622. https://doi.org/10.1371/journal.pone.0021622

4. Argyros O, Wong S, Niceta M, et al (2008) Persistent episomal transgene expression in liver following delivery of a scaffold/matrix attachment region containing non-viral vector. Gene Ther 13

5. Argyros O, Wong SP, Fedonidis C, et al (2011) Development of S/MAR minicircles for enhanced and persistent transgene expression in the mouse liver. J Mol Med 89:515–529. https://doi.org/10.1007/s00109-010-0713-3

6. Hagedorn C, Schnödt-Fuchs M, Boehme P, et al (2017) S/MAR Element Facilitates Episomal Long-Term Persistence of Adeno-Associated Virus Vector Genomes in Proliferating Cells. Hum Gene Ther 28:1169–1179. https://doi.org/10.1089/hum.2017.025

7. Wang X-Y, Zhang X, Wang T-Y, et al (2019) Shortened nuclear matrix attachment regions are sufficient for replication and maintenance of episomes in mammalian cells. Mol Biol Cell 30:2761–2770. https://doi.org/10.1091/mbc.E19-02-0108

8. Prevent Nucleic Acid Degradation Before Extraction. In: ZYMO Res. https://www.zymoresearch.com/blogs/blog/prevent-nucleic-acid-degradation. Accessed 3 Sep 2021

9. Marillonnet S, Grützner R (2020) Synthetic DNA Assembly Using Golden Gate Cloning and the Hierarchical Modular Cloning Pipeline. Curr Protoc Mol Biol 130:e115. https://doi.org/10.1002/cpmb.115

10. Kozak M (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 44:283–292. https://doi.org/10.1016/0092-8674(86)90762-2

11. Edelheit O, Hanukoglu A, Hanukoglu I (2009) Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. BMC Biotechnol 9:61. https://doi.org/10.1186/1472-6750-9-61

12. Rong M, He B, McAllister WT, Durbin RK (1998) Promoter specificity determinants of T7 RNA polymerase. Proc Natl Acad Sci 95:515–519. https://doi.org/10.1073/pnas.95.2.515